



MAR 09 2004

1/23

BAL human cDNA

5' UTR:

GGGCTTCGTGTTCCCTGGGTGCTGACCGTGCACTCCCCGCCGCCGAGGACTTAGAGCTCTGGAAGT
AGCTCTCCAGCTTCCTTCGTACTCGGGGGCCGACTTGTACACCCGCACGAGGAGCGGGGACGGC
GGGCGCAGAAGTGGGCCACCATATCTGGAAATACAGTCTATGCTTTGAAGCGCAAAAGGGAATA
AACATTTAAAGACTCCCCCGGGGACCTGGAGG

Coding: alternatively spliced sequence in bold characters

ATGGACTTTTCCATGGTGGCCGGAGCAGCAGCTTACAATGAAAAATCAGGTAGGATTACCTCGCT
CTCACTCTTGTTCAGAAAGTCTTTGCTCAGATCTTTCCTCAGTGGAGAAAGGGGAATACAG
AAGAATGTCTCCCCTACAAGTGCTCAGAGACTGGTGCTCTTGGAGAAAATATAGTTGGCAAAAT
TCCCATTAACCACAATGACTTCAAAATTTAAAAAATAATGAGCGTCAGCTGTGTGAAGTCTCCA
GAATAAGTTGGCTGTATCTCTACCCTGGTCTCTCCAGTTCAGGAAGGCAACAGCAAATCTCTGCA
AGTGTTTCAAGAAAATGCTGACTCCTAGGATAGAGTTATCAGTCTGGAAAGATGACCTCACCACAC
ATGCTGTTGATGCTGTGGTGAATGCAGCCAATGAAGATCTTCTGCAATGGGGGAGGCCTGGCCCTGG
CCCTGGTAAAGCTGGTGGATTTGAAATCCAAGAAGAGAGCAAACAGTTTGTGGCAGATATGGT
AAAGTGTCAGCTGGTGAGATAGCTGTACGGGAGCAGGAGGCTTCCCTGCAAACAGATCATCCA
TGCTGTTGGGCCTCGGTGGATGGGAATGGGATAAACAGGGATGTACTGGAAAGCTGCAGAGGGCCA
TTGTAAGTATTCTGAATTATGTCATCTATAAAATACTCACATTAAGACAGTAGCAATTCCAGCCT
TGAGCTCTGGGATTTTTCAGTTCCCTCTGAATTTGTGTACAAAGACTATTGTAGAGACTATCCGGGT
TAGTTTGAAGGAAGCCAATGATGAGTAATTTGAAAGAAATTCACCTGGTGAGCAATGAGGACC
CTACTGTTGCTGCCTTTAAAGCTGCTTCAGAATTCATCCTAGGGAAGAGTGAGCTGGGACAAGAAA
CCACCCCTTCTTTCAATGCAATGGTCGTGAACAACCTGACCCTCCAGATTGTCCAGGGCCACATTG
AATGGCAGACGGCAGATGTAATTGTAAATCTGTAAACCCACATGATATTACAGTTGGACCTGTGG
CAAAGTCAATTCTACAACAAGCAGGAGTTGAAATGAAATCGGAATTTCTTGCCACAAAGGCTAAA
CAGTTTCAACGGTCCCAGTTGGTACTGGTCACAAAAGGATTTAACTTGTCTGTAAATATATATAC
CATGTACTGTGGCATTGAGAATTTCCCTAAACCTCAGATATTAAACATGCAATGAAGGAGTGTG
GAAAAATGCATTGAGCAAAATATACTTCCATTTCCCTTCCCTGCCCCTGGGACTGGAAACATGGAA
ATAAAGAAGGAAACAGCAGCAGAGATTTTGTGTTGATGAAGTTTAAACATTTGCCAAAGACCATGT
AAAACACCAGTTAACTGTAAAATTTGTGATCTTTCCAACAGATTGAGAGATATATAAGGCTTTTCAG
TTCTGAAATGGCAAAGAGGTCCAAGATGCTGAGTTTGAACAATTACAGTGTCCCCCAGTCAACCA
GAGAGGAGAAAAGAGAAAATGGGCTTGAAGCTAGATCTCCTGCCATCAATCTGATGGGATTCAAC
GTGGAAGAGATGTAGTGAGGCCCACGCATGGATCCAAAGAATCCTGAGTCTCCAGAACCACCACA
TCATTGAGAATAATCATATTCTGTACCTTGGGAGAAAGGAACATGACATTTTGTCTCAGCTTCAGA
AACTTCAAGTGTCTCCATCACAGAAATTATCAGCCCAGGAAGGACAGAGTTAGAGATTGAAGGA
GCCCCGGCTGACCTCATTGAGGTGGTTATGAACATTGAAGATATGCTTTGTAAAGTACAGGAGGA
AATGGCAAGGAAAAAGGAGCGAGGCCTTTGGCGCTCGTTAGGACAGTGGACTATTTCAGCAACAAA
AAACCCAAGACGAAATGAAAGAAAATATCATATTTCTGAAATGTCTGTGCCTCCAACCTCAAGAG
CTTCTAGATCAAAAGAAACAGTTTGAAAAATGTGGTTTGCAGGTTCTAAAGGTGGAGAAGATAGA
CAATGAGGTCTTATGGCTGCCTTTCAAAGAAAAGAAAAATGATGGAAGAAAAACTGCACAGGC
AACCTGTGAGCCATAGGCTGTTTCAGCAAGTCCCATACCAGTTCTGCAATGTGGTATGCAGAGTTG
GCTTTCAAAGAATGTACTCGACACCTTGCGATCCAAAATACGGAGCTGGCATATACTTCACCAAGA
ACCTCAAAAACCTGGCAGAGAAGGCCAAGAAAAATCTCTGCTGCAGATAAGCTGATCTATGTGTTT
GAGGCTGAAGTACTCACAGGCTTCTTCTGCCAGGGACATCCGTAAATATTGTTCCCCCACCCTG
AGTCCTGGAGCTATAGATGGTCATGACAGTGTGGTTGACAAATGTCTCCAGCCCTGAAACCTTTGTT
ATTTTTAGTGGCATGCAGGCTATACCTCAGTATTTGTGGACATGCACCCAGGAATATGTACAGTCA
CAAGATTACTCATCAGGACCAATGAGACCTTTGCACAGCATCCTTGGAGGGGATTGCAAGTGG
CAGCCCTGTTGATTAA

Fig. 1



2/23

3' UTR

TCTCTACATCATTTTAAACAGCTGGTATGGCCTTACCTTGGGTGAACTAACCAAATAATGACCATCG
ATGGCTCAAAGAGTGGCTTGAATATATCCCATGGGTATCTGTATGGACTGACTGGGTATTGAAA
GGACTAGCCACATACTAGCATCTTAGTGCCTTTATCTGTCTTTATGTCTTGGGGTTGGGGTAGGTAG
ATACCAAATGAAACACTTTCAGGACCTTCCTTCCTCTTGCAGTTGTTCTTTAATCTCCTTTACTAGA
GGAGATAAAATATTTTGCATATAATGAAGAAATTTTCTAGTATATAACGCAGGCCTTTTATTTTCTA
AAATGATGATAGTATAAAAATGTTAGGATAACAGAATGATTTTAGATTTTCCAGAGAATATTATAA
AGTGCTTTAGGTATGAAAATAAATCATCTTTGTCTGATTAAAAAAAAAAAAA

BAL human protein: alternatively spliced (Bold characters)

MDFSMVAGAAAYNEKSGRITSL**SL**L**FQKVFAQIFPQWRKGNTEECLPYKC**SETGALGENYSW
QIPINHNDFKILKNNERQLCEVLQNKFGCISTLVSPVQEGNSKSLQVFRKMLTPRIELSVWKDDLTT
HAVDAVVNAANEDLLHGGGLALALVKAGGFEIQEESKQFVARYGKVSAGEIAVTGAGRLPCKQIIHAV
GPRWMEWDKQGCTGKLQRAIVSILNYVIYKNTHIKTVAIPALSSGIFQFPLNLCTKTIVETIRVSLQGKP
MMSNLKEIHLVSNEDPTVAAFKAASEFILGKSELGQETTPSFNAMVVNNLTQIVQGHIEWQTADVIVN
SVNPHDITVGPVAKSILQQAGVEMKSEFLATKAKQFQRSQVLVTKGFNLFCYIYHVLWHSEFPKPQI
LKHAMKECLEKCIEQNITSISFPALGTGNMEIKKETAAEILFDEVLTFAKD
HVKHQLTVKFVIFPTDLEIYKAFSSEMAKRSKMLSLNNYSVPQSTREEKRENGLEARSPAINLMGFNVE
EMYEAHAWIQRILSLQNHIIENNHIYLGKHEHDILSQLQKTSSVSITEIISPGRTELEIEGARADLIEVV
MNIEDMLCKVQEEEMARKKERGLWRSLGQWTIQQKQKTQDEMKENIIFLKCPVPPTQELLDQKKQFEKC
GLQVLKVEKIDNEVLMAAFQRRKKMMEEKLHRQPVSHRLFQQVPYQFCNVVCRVGFQRMYSTPCDP
KYGAGIYFTKNLKNLAEKAKKISAADKLIYVFEAEVLTGFFCQGHPLNIVPPPLSPGAIDGHDSVVDNVS
SPETFVIFSGMQAIPQYLWTCTQEYVQSQDYSSGPMRPFQHPWRGFASGSPVD

Fig. 1A

MAR 09 2004



3/23

BAL mouse cDNA:

MAR 03 2004

5' UTR

AGGAACGGAAGTTTGGCGGGAACCCGGATTCCCAGGTTCAAGGGTGGAGCGGAATA
GAGGGAAACAGGCCACCATCTCCTCGATCTACAGACTACACTTGGAAACACAAACAAATATAAAT
ATCTGAAGACCCACGTGGGACCTGAAGAATGGCCTATTAC

Coding region (shorter form only)

ATGGATACATGGGCGGCAGTCCCCGCCGAAAGACCAGCCAACAATTCTCTTGAAGAACATTATAG
ATGGCAAATTTCCATTAAACACAATGTCTTCGAAATTTTAAAGAGCAATGAGAGTCAGCTATGTGA
AGTCCCTCCAAAATAAGTTTGGATGCATCTCTACCCTGAGCTGTCCAACCTCTAGCAGGGAGCAGCTC
TCCTGCTCAGAGAGTCTTCAGAAGGACCCTGATCCCTGGGATAGAGTTATCTGTCTGGAAGGATGA
CCTTACCAGACACGTTGTTGATGCTGTGGTGAACGCAGCCAATGAAAACCTTTTGCATGGAAGTGG
CCTGGCCCGGAAGCTTGGTGAAGAACTGGTGGCTTTGAAATCCAAGAAGAGAGCAAAAAGAATCATTG
CCAACGTTGGTAAAATCTCAGTTGGTGAATCGCTATCACCGGTGCGGGGAGACTTCCTTGCCATT
TGATTATCCATGCGGTTGGACCTCGGTGGACAGTTACGAACAGCCAGACAGCTATCGAATTACTGA
AATTTGCCATTAGGAACATTCTAGATTATGTCACCAAATATGATCTACGCATTAAGACAGTAGCAA
TTCCAGCCCTGAGCTCTGGAATTTTCCAGTTCCCTCTGGATTTGTGTACAAGCATAATTTTAGAAAC
TATCCGGCTTTATTTCCAAGACAAGCAAATGTTCCGTAATTTGAGAGAGATTCTATCTGGTGAGCAA
TGAGGACCCCACTGTTGCGTCTTTAAATCCGCCTCAGAAAGCATCCTAGGGAGGGACCTGAGCTC
TTGGGGGGGTCCAGAACTGACCCTGCTTCCACCATGACTCTTCGCATCGGCCGGGCGCTGACTCT
CCAGATTGTCCAAGGCTGTATTGAAATGCAAACAACAGATGTAATTGGTAATTCTGGATACATGCA
GGATTTTAAATCAGGACGAGTGGCACAGTCGATTCTTAGACAAGCAGGGGTTGAAATGGAAAAG
AACTTGACAAGGTTAACCTGTCCACAGATTATCAAGAGGTGTGGGTACAAAAGGATTTAAATGT
CCTGTCTAGTATGTCTTCCATGTGGCATGGCATTCCCAAATCAACAAATACCAGATATTGAAAGATG
CAATGAAGTCTGTCTAGAAAAATGCCTTAAACCAGATATAAATTCCATTTCTCTTCTGCTCTCG
GGACAGGATTGATGGATTTGAAGAAGAGTACAGCAGCTCAGATAATGTTTGAGGAAGTTTTTGCA
TTTGCTAAAGAGCACAAGGAAAAACGCTAACTGTAAAGATTGTGATCTTTCCAGTAGATGTGGA
GACGTACAAGATTTTTTATGCTGAAATGACAAAAAGGTCCAACGAGCTGAATCTCAGCGGTAATA
GTGGTGCTTTAGCCCTGCAGTGGTCCAGTGGGGAGCAAAGAAGAGGCGGCCTTGAAGCTGGATCT
CCTGCCATCAATCTCATGGGTGTAAAGTGGGAGAGATGTGTGAGGCCAGGAATGGATTGAAAG
GTTGCTGGTCTCCCTGGACCACCACATCATTGAGAATAATCATATTCTCTATCTTGGGAAAAAAGA
GCACGACGTGCTGTCTGAGCTCCAGACCAGCACAAAGAGTCTCCATTTAGAGACTGTCAGTCCAA
GAACGGCCACTTTGGAGATTAAAGGTCCCCAGGCTGACCTCATTGACGCAGTTATGAGGATTGAAT
GTATGCTGTGTGACGTTTCAAGGAAGAAGTGGCAGGAAAAAGGGAGAAAAATCTTTGGAGCTTGTC
GGACAGGGGACCAACCAGCAAGAAAACTGGATAAAATGGAAGAATCGTACACATTTCAACGAT
ACCCAGCATCATTAACCTCAGGAACCTCAGGACCGAAAGAAACAGTTTGAAGAGTGTGGCTTGTGG
GTTGTGTCAGGTGGAGCAGATAGACAATAAGGTGCTGCTGGCTGCCTTCCAAGAGAAGAAGAAAT
GATGGAAGAGAGGACGCCAAAGGGATCTGGGAGCCAAAGGTTGTTTCAGCAGGTCCCACATCAGT
TCTGCAATACGGTGTGACAGTCCGCTTCCACAGAATGTATTGACATCCTATAACCCAGTTTATG
GAGCCGGCATATATTTACCAAGAGCCTCAAAAATCTAGCAGACAAGGTCAAGAAAACCTCAAGC
ACAGACAAGCTAATCTATGTGTTTGAAGCAGAAGTACTCACAGGGTCTTCTGTGAGGGTAATTC
TCAAATATCATCCCTCCACCATTGAGTCTGGGGCCTTAGATGTCAATGACAGCGTAGTTGACAAAT
GTTTCCAGCCCTGAAACCATTGTTGTTTTTAATGGCATGCAGGCCATGCCCCTGTACTTGTGGACTT
GCACACAGGATAGGACATTCTCACAGCATCCGATGTGGTCACAGGACTACTCATCAGGACCAGGA
ATGGTCTCTTCGCTGCAGTCCCTGGGAATGGGTCTTAAATGGCAGCTCTGTTTAG

3' UTR:

TGTCTACATCAGTTTAAACAAGCAGAAGGGGTTGAGAGAACTGACAAAATGATAAATAACAGGTTA
CCTGTTTCAAGATGATGGGGTCACTAAAGGCACCGACCACACACTAGCATCATAGTGCCTTTGTCTT
TACCTCTGGGCTTGACTGGGCAGATGCCAGCTAAACTTCCTCACTGTCTT

Fig. 2



MAR 09 2004

4/23

TTCTATTTGATATCTTTTCATCTCCTTTTCCTATAGGTGACAGCAAGAATACTTTATATAGAACAAGGA
TATTTTTTTTCAAGCCTGTTATTTTCTAAAATGATAGCACAACTAGGACAACAGGATGATTTTCAGG
TTTTCTATATAATTTATAAAGTGCTTTGGATATCCAAATAAATCACCTTTGTCTGAGT

BAL mouse protein (shorter form):

MDTWAAAPAERPANNLEEYRWQIPIKHNVFEILKSNESQLCEVLQNKFGCISTLSCPTLAGSSS
PAQRVFRRTLIPGIELSVWKDDLTRHVVDVAVNAANENLLHGSLAGSLVKTGGFEIQEESKRIIA
NVGKISVGGAITGAGRLPCHLIHAGVPRWTVTNSQTAEILLKFAIRNILDYVTKYDLRIKTVAIPA
LSSGIFQFPLDLCTSIILETIRLYFQDKQMFGNLREIHLVSNEDPTVASFKSASESILGRDLSSWGGP
ETDPASTMTLRIGRLTLQIVQGCIEMQTTDVIGNSGYMQDFKSGRVAQSILRQAGVEMEKELDK
VNLSTDYQEVWVTKGFKLSCQYVFHVAWHSQINKYQILKDAMKSCLEKCLKPDINSISFPALGT
GLMDLKKSTAAQIMFEEVFATAKEHKEKTLTVKIVIFPVDVETYKIFYAEMTKRSNELNLSGNSG
ALALQWSSGEQRRGGLEAGSPA INLMGVKVGEMCEAQEWIERLLVSLDHIIENNHILYLKKE
HDVLSELQTSTRVSISSETVSPRTATLEIKGPQADLIDAVMRIECMLCDVQEEVAGKREKNLWSLS
GQGTNQOEKLDKMEESYTFQRYPASLTQELQDRKKQFEKCLWVQVEQIDNKVLLAAFQE K
KKMMEERTPKGSGSQRLFQQVPHQFCNTVCRVGFHRMYSTSYNPVYGAGIYFTKSLKNLADKV
KKTSSTDKLIYVFEEAEVLTGSFCQGNSSNIIPPLSPGALDVNDSVVDNVSSPETIVVFNGMQAMP
LYLWTCTQDRFTSQHPMWSQDYSSGPGMVSSLQSEWVLNGSSV

Fig. 2A



5/23

>_ BAL Human protein 819 aa vs.
>_ BAL Mouse protein 826 aa
scoring matrix: , gap penalties: -12/-2
61.5% identity; Global alignment score: 3158

```

      10      20      30      40      50      60
610015 MDFSMVAGAAAYNEKSETGALGENYSWQIPINHNDFKILKNNERQLCEVLQNKFGCISTL
      :: . : : : . : : : : : : : : : : : : : : : : : : : : : : : : :
MD----TWAAAPAERPANNSSLEEHYRWQIPIKHNVFEILKSNESQLCEVLQNKFGCISTL
      10      20      30      40      50

      70      80      90     100     110
610015 VSPVQEGNSKSLQ-VFRKMLTPRIELSVWKDDLTHAVDAVVNAANEDLLHGGGLALALV
      :. :. : : : : : : : : : : : : : : : : : : : : : : : : : : :
SCPTLAGSSSPAQRVFRRTLIPGIELSVWKDDLTRHVVDVVNAANENLLHGSGLAGSLV
      60      70      80      90     100     110

     120     130     140     150     160     170
610015 KAGGFEIQEESKQFVARYGKVSAGEIAVTGAGRLPCKQIIHAVGPRWMEWDKQGCTGKLQ
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
KTGGFEIQEESKRIIANVGKISVGGIAITGAGRLPCHLIIHAVGPRWTVTVNSQTAEILLK
      120     130     140     150     160     170

     180     190     200     210     220     230
610015 RAIVSILNYVIYKNTHIKTVAIPALSSGIFQFPLNLCTKTIVETIRVSLQGKPMMSNLKE
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
FAIRNILDYVTKYDLRIKTVAIPALSSGIFQFPLDLCTSIILETIRLYFQDKQMFGNLR
      180     190     200     210     220     230

     240     250     260     270     280     290
610015 IHLVSNEEDPTVAAFKAASEFILGK---SELGQETTP--SFNAMVVNNLTQIVQGHIEWQ
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
IHLVSNEEDPTVASFKSASESILGRDLSSWGGPETDPASTMTLRIGRGLTLQIVQGCIEMQ
      240     250     260     270     280     290

     300     310     320     330     340     350
610015 TADVIVNSVNPHTITVGPVAKSILQQAGVEMKSEFLATKAKQFQRSQVLVTKGFNLFCCK
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
TTDVIGNSGYMQDFKSGRVAQSILRQAGVEMEKEL--DKVNLSTDYQEVWVTKGFKLSCQ
      300     310     320     330     340     350

     360     370     380     390     400     410
610015 YIIHVLWHSEFPKPKQILKHAMKECLEKQIEQNITSISFPALGTGNMEIKKETAAEILFDE
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
YVFHVAWHSQINKYQILKDAMKSCLEKCLKPDINSISFPALGTGLMDLKKSTAAQIMFEE
      360     370     380     390     400     410

     420     430     440     450     460     470
610015 VLTFAKDHVKHQLTVKFVIFPTDLEIYKAFSSEMAKRSKMLSLNNYS---VPQSTREEKR
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
VFATAKEHKEKTLTVKIVIFPVDVETYKIFYAEMTKRSNELNLSGNSGALALQWSSGEQR
      420     430     440     450     460     470
```

Fig. 3

6/23

	480	490	500	510	520	530
610015	ENGLE	ARSPAINLMGFN	VEEMYEAHAWIQ	RILSLQNHIIENN	HILYLGRKEHD	ILSQLQ

	RGGL	EAGSPAINLMGV	KVGEMCEAQEWI	ERLLVSLDHHII	ENNHLILGKKEH	DVLSLQ
	480	490	500	510	520	530
	540	550	560	570	580	590
610015	KTSSVS	SITEIISGRTELE	IEGARADLIEVVM	NIEDMLCKVQEEM	ARKKERGLWRS	LQGW

	TSTRV	SISSETVSPRTAT	LEIKGPQADLID	AVMRIECMLCDV	QEEVAGKREKNL	WSLSGQ
	540	550	560	570	580	590
	600	610	620	630	640	650
610015	TIQQQKTQ	DEMKENIIFLKCP	VPPTQELLDQKKQ	FEKCGLQVLKVE	KIDNEVLMAAF	QORK
	:
	TNQ	QEKLDKMEESYTF	QRYPASLTQELQ	DRKKQFEKCGLW	VVQVEQIDNKVLL	AAFAQEK
	600	610	620	630	640	650
	660	670	680	690	700	710
610015	KKMMEEKL	HRQPVSHRLFQQ	VPYQFCNVVCRV	GFQRMYSTPCDP	KYGAGIYFTKNL	KNLA

	KKMME	ERTPKGSGSQRL	FQQVPHQFCNTV	CRVGFHRMYS	TSYNPVYGAGI	YFTKSLKNLA
	660	670	680	690	700	710
	720	730	740	750	760	770
610015	EKAKKISA	ADKLIYVF	EAEVLTGFFCQ	GHPLNIVPPPL	SPGAIDGHDS	VVDNVSSPETFV

	DKVK	KTSSTDKLIYV	F	EAEVLTGSFCQ	GNSNIIPPLSP	GALDVNDVSSPETIV
	720	730	740	750	760	770
	780	790	800	810		
610015	IFSGM	QAIPQYLWTCTQ	E--YVQ---	SQDYSSGPMRPF	AQHPWRGFASG	SPVD

	VFNGM	QAMPLYLWTCTQ	DRTF	SQHPMWSQDYSSG	PGMVSSLQSW	EVLNGSSV-
	780	790	800	810	820	

Fig. 3A



7/23

Comparison of:

(A) 7486572155.52.67.361 > BAL Human

3244n

(B) 7486572155.52.67.362 > BAL Mouse

3024n

using matrix file: DNA, gap penalties: -16/-4

71.7% identity in 2916 nt overlap; score: 5444

```

      370      380      390      400      410      420
- CCCTACAAGTGCTCAGAGACTGGTGCTCTTGGAGAAACTATAGTTGGCAAATTTCCCAT
  :: : : : :: : : ::::: :::: : ::::: ::::: ::::: :::::
- CCCGCCGAAAGACCAGCCAACAATTCTCTTGAAGAACATTATAGATGGCAAATTTCCCAT
      200      210      220      230      240      250

      430      440      450      460      470      480
- AACCACAATGACTTCAAATTTTAAAAAATAATGAGCGTCAGCTGTGTGAAGTCTCCAG
  :: ::::: ::::: ::::: ::::: : ::::: ::::: ::::: ::::: :::::
- AAACACAATGTCTTCGAAATTTTAAAGAGCAATGAGAGTCAGCTATGTGAAGTCTCCAA
      260      270      280      290      300      310

      490      500      510      520      530      540
- AATAAGTTTGGCTGTATCTCTACCCTGGTCTCTCCAGTTCAGGAAGGCAACAGCAAATCT
  ::::: ::::: ::::: ::::: ::::: ::::: : ::::: ::::: ::::: :::::
- AATAAGTTTGGATGCATCTCTACCCTGAGCTGTCCAACCTCTAGCAGGGAGCAGCTCTCCT
      320      330      340      350      360      370

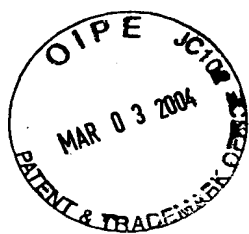
      550      560      570      580      590      600
- CTGCA---AGTGTTTCAGAAAAATGCTGACTCCTAGGATAGAGTTATCAGTCTGGAAGAT
  :: ::::: ::::: : ::::: ::::: ::::: ::::: ::::: ::::: :::::
- GCTCAGAGAGTCTTCAGAAGGACCCTGATCCCTGGGATAGAGTTATCTGTCTGGAAGGAT
      380      390      400      410      420      430

      610      620      630      640      650      660
- GACCTCACCACACATGCTGTTGATGCTGTGGTGAATGCAGCCAATGAAGATCTTCTGCAT
  ::::: ::::: ::::: : ::::: ::::: ::::: ::::: ::::: ::::: :::::
- GACCTTACCAGACACGTTGTTGATGCTGTGGTGAACGCAGCCAATGAAAACCTTTTGCAT
      440      450      460      470      480      490

      670      680      690      700      710      720
- GGGGGAGGCCTGGCCCTGGCCCTGGTAAAAGCTGGTGGATTGAAATCCAAGAAGAGAGC
  :: : ::::: : ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
- GGAAGTGGCCTGGCCGGAACCTTGGTGAAACTGGTGGCTTTGAAATCCAAGAAGAGAGC
      500      510      520      530      540      550

      730      740      750      760      770      780
- AAACAGTTTGTGGCCAGATATGGTAAAGTGTCAGCTGGTGAGATAGCTGTCACGGGAGCA
  :: : ::::: ::::: : ::::: ::::: ::::: ::::: ::::: ::::: :::::
- AAAAGAATCATTGCCAACGTTGGTAAAATCTCAGTTGGTGGAATCGCTATCACCGGTGCG
      560      570      580      590      600      610
```

Fig. 4



8/23

790 800 810 820 830 840
GGGAGGGCTTCCTTGCAAACAGATCATCCATGCTGTTGGGCCTCGGTGGATGGAATGGGAT
:::
GGGAGACTTCCCTTGCCATTGATTATCCATGCGGTTGGACCTCGGTGGACAGTTACGAAC
620 630 640 650 660 670

850 860 870 880 890 900
AAACAGGGATGTACTGGAAAAGCTGCAGAGGGCCATTGTAAGTATTCTGAATTATGTCATC
:
AGCCAGACAGCTATCGAATTACTGAAATTTGCCATTAGGAACATTCTAGATTATGTCACC
680 690 700 710 720 730

910 920 930 940 950 960
TATAAAAACTCACATTAAGACAGTAGCAATTCCAGCCTTGAGCTCTGGGATTTTTTCAG
:
AAATATGATCTACGCATTAAGACAGTAGCAATTCCAGCCCTGAGCTCTGGAATTTTTCCAG
740 750 760 770 780 790

970 980 990 1000 1010 1020
TTCCTCTGAATTTGTGTACAAAGACTATTGTAGAGACTATCCGGGTTAGTTTGCAAGGG
:
TTCCTCTGGATTTGTGTACAAGCATAATTTAGAACTATCCGGCTTTATTTCCAAGAC
800 810 820 830 840 850

1030 1040 1050 1060 1070 1080
AAGCCAATGATGAGTAATTTGAAAGAAATTCACCTGGTGAGCAATGAGGACCCTACTGTT
:
AAGCAAATGTTTCGGTAATTTGAGAGAGATTTCATCTGGTGAGCAATGAGGACCCCCTGTT
860 870 880 890 900 910

1090 1100 1110 1120 1130
GCTGCCTTTAAAGCTGCTTCAGAATTCATCCTAGGGAAGAG--TGAGCT-----GGGA
:
GCGTCCTTTAAATCCGCCTCAGAAAGCATCCTAGGGAGGGACCTGAGCTCTTGGGGGGGT
920 930 940 950 960 970

1140 1150 1160 1170 1180
CAAGAAACCACCCCTTCTTTCA--ATG-CAATGGTCGTGAACAA---CCTGACCCTCCAG
:
CCAGAAACTGACCCTGCTTCCACCATGACTCTTCGCATCGGCCGGGGCCTGACTCTCCAG
980 990 1000 1010 1020 1030

1190 1200 1210 1220 1230 1240
ATTGTCCAGGGCCACATTGAATGGCAGACGGCAGATGTAATTGTTAATTCTGTAAACCCA
:
ATTGTCCAAGGCTGTATTGAAATGCAAACAACAGATGTAATTGGTAATTCTGGATACATG
1040 1050 1060 1070 1080 1090

Fig. 4A



9/23

1250 1260 1270 1280 1290 1300
CATGATATTACAGTTGGACCTGTGGCAAAGTCAATTCTACAACAAGCAGGAGTTGAAATG
:: :: :: : :::: :::::: :::: :::::: :::::::::: ::::::::::
CAGGATTTTTAAATCAGGACGAGTGGCACAGTCGATTCTTAGACAAGCAGGGGTTGAAATG
1100 1110 1120 1130 1140 1150

1310 1320 1330 1340 1350 1360
AAATCGGAATTTCTTGCCACAAAGGCTAAACAGTTTCAACGGTCCCAGTTGGTACTGGTTC
:: :::: :::: :: : :: :: : : : : : : : : : : : : : : : : : :
GAAAAGGAA---CTTGACA--AGGTTAACCTGTCCACAGATTATCAAGAGGTGTGGGTTC
1160 1170 1180 1190 1200

1370 1380 1390 1400 1410 1420
ACAAAAGGATTTAACTTGTTCTGTAAATATATATACCATGTACTGTGGCATTTCAGAATTT
:::::::::::: :::: ::::
ACAAAAGGATTTAAATTGTCCTGTCTCAGTATGTCTTCCATGTGGCATGGCATTCCCAAATC
1210 1220 1230 1240 1250 1260

1430 1440 1450 1460 1470 1480
CCTAAACCTCAGATATTA AACATGCAATGAAGGAGTGT TGGAAAAATGCATTGAGCAA
:: : :::: :: : :::: ::::
AACAATAACCAGATATTGAAAGATGCAATGAAGTCCTGTCTAGAAAAATGCCTTAAACCA
1270 1280 1290 1300 1310 1320

1490 1500 1510 1520 1530 1540
AATATAACTTCCATTT CCTTTCTG CCGCTTGGGACTGG AAACATGG AAAATAAGAAGGAA
:::::: :::::::::::::: :: : :::: :: : :::: : : :::: : : :::::
GATATAAATTCCATTT CCTTTCTG CTCTCGGGACAGGATTG ATGGATTG AAGAAGAGT
1330 1340 1350 1360 1370 1380

1550 1560 1570 1580 1590 1600
ACAGCAGCAGAGATTTT GTTTGATGA AGTTTTTA ACATTTGCC AAAGACCATGT AAAACAC
:::::::: :: : :::: :::: :::: :::: :::: :::: : : : : : : : : : :
ACAGCAGCTCAGATAAT GTTTGAGGAAGTTTTT GCATTTGCT AAAGAGCACA AGGAAAAA
1390 1400 1410 1420 1430 1440

1610 1620 1630 1640 1650 1660
CAGTTAACTGTAAATTT GTGATCTTTCCA ACAGATTTGG AGATATATA AGGCTTTTCAGT
: :::::::::: :::: :::: :::: :::: :::: :::: : : : : : : : : : :
ACGCTAACTGTAAAGATT GTGATCTTTCCAG TAGATGTGG AGACGTACA AGATTTTTTTAT
1450 1460 1470 1480 1490 1500

1670 1680 1690 1700 1710
TCTGAAATGGCAAAGAGGTCCAAGATGCTGAGTTTGAACAATTACAGTGT-----C
:::::::: : : : :::: ::::
GCTGAAATGACAAAAGGTCCAACGAGCTGAATCTCAGCGGTAATAGTGGTGCTTTAGCC
1510 1520 1530 1540 1550 1560

Fig. 4B



11/23

2200 2210 2220 2230 2240 2250
AAAGAAACAGTTTGAAAAATGTGGTTTGCAGGTTCTAAAGGTGGAGAAGATAGACAATGA
:::
AAAGAAACAGTTTGAAAAGTGTGGCTTGTGGGTTGTGCAGGTGGAGCAGATAGACAATAA
2050 2060 2070 2080 2090 2100

2260 2270 2280 2290 2300 2310
GGTCCTTATGGCTGCCTTTCAAAGAAAGAAGAAAATGATGGAAGAAAAACTGCACAGGCCA
::: ::
GGTGTCTGCTGGCTGCCTTCCAAGAGAAGAAGAAAATGATGGAAGAGAGGACGCCAAAGGG
2110 2120 2130 2140 2150 2160

2320 2330 2340 2350 2360 2370
ACCTGTGAGCCATAGGCTGTTTCAGCAAGTCCCATAACCAGTTCTGCAATGTGGTATGCAG
:
ATCTGGGAGCCAAAGGTTGTTTCAGCAGGTCCCACATCAGTTCTGCAATACGGTGTGCAG
2170 2180 2190 2200 2210 2220

2380 2390 2400 2410 2420 2430
AGTTGGCTTTCAAAGAATGTACTCGACACCTTGCGATCCAAAATACGGAGCTGGCATATA
:::
AGTCGGCTTCCACAGAATGTATTTCGACATCCTATAACCCAGTTTATGGAGCCGGCATATA
2230 2240 2250 2260 2270 2280

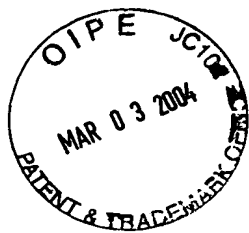
2440 2450 2460 2470 2480 2490
CTTCACCAAGAACCTCAAAAACCTGGCAGAGAAGGCCAAGAAAATCTCTGCTGCAGATAA
:
TTTCACCAAGAGCCTCAAAAATCTAGCAGACAAGGTCAAGAAAACCTCAAGCACAGACAA
2290 2300 2310 2320 2330 2340

2500 2510 2520 2530 2540 2550
GCTGATCTATGTGTTTGAGGCTGAAGTACTCACAGGCTTCTTCTGCCAGGGACATCCGTT
:::
GCTAATCTATGTGTTTGAGGCAGAAGTACTCACAGGGTCCTTCTGTGTCAGGGTAATTCCTC
2350 2360 2370 2380 2390 2400

2560 2570 2580 2590 2600 2610
AAATATTGTTCCCCCACCCTGAGTCCTGGAGCTATAGATGGTCATGACAGTGTGGTTGA
:
AAATATCATCCCTCCACCATTGAGTCCTGGGGCCTTAGATGTCAATGACAGCGTAGTTGA
2410 2420 2430 2440 2450 2460

2620 2630 2640 2650 2660 2670
CAATGTCTCCAGCCCTGAAACCTTTGTTATTTTTAGTGGCATGCAGGCTATACCTCAGTA
:
CAATGTTTCCAGCCCTGAAACCATTTGTTGTTTTAATGGCATGCAGGCCATGCCCTGTA
2470 2480 2490 2500 2510 2520

Fig. 4D



12/23

```
2680      2690      2700      2710      2720
-   TTTGTGGACATGCACCCAGGA----ATATGTACA-----GTCACAAGATTA
-   ::::: : : : : : : : : : : : : : : : : : : : : : : : : :
-   CTTGTGGACTTGACACAGGATAGGACATTCTCACAGCATCCGATGTGGTCACAGGACTA
-   2530      2540      2550      2560      2570      2580

      2730      2740      2750      2760      2770      2780
-   CTCATCAGGACCAATGAGACCCTTTGCACAGCATCCTTGGAGGGGATTTCGAAAGTGGCAG
-   ::::: : : : : : : : : : : : : : : : : : : : : : : : : :
-   CTCATCAGGACCAGGAATGGTCTCTTCGCTGCAGTCCTGGGAATGGGTCTTAAATGGCAG
-   2590      2600      2610      2620      2630      2640

      2790      2800      2810      2820      2830      2840
-   CCCTGTTGATTAATCTCTACATCATTTTAAACAGCTGGTATGGCCTTACCTTGGGTGAAC
-   : : : : : : : : : : : : : : : : : : : : : : : : : :
-   CTCTGT---TTAGTGTCTACATCAGTTTAAACAAGCAGAAAGGG-----TTGAGAGAACT
-   2650      2660      2670      2680      2690

      2850      2860      2870      2880      2890      2900
-   AACCAAATAATGACCATCGATGGCTCAAAGAGTGGCTTGAATATATCCCATGGGTATCT
-   : : : : : : : : : : : : : : : : : : : : : : : : : :
-   GACAAAATGAT-----AAATA-----ACAGGTTACCT
-   2700      2710

      2910      2920      2930      2940      2950      2960
-   GTATGGACTGACTGGGTTATTGAAAGGACTAGCCACATACTAGCATCTTAGTGCCTTTAT
-   : : : : : : : : : : : : : : : : : : : : : : : : : :
-   GTTCAGAATGATGGGGTCACTAAAGGCACCGACCACACACTAGCATCATAGTGCCTTT--
2720      2730      2740      2750      2760      2770

      2970      2980      2990      3000      3010      3020
-   CTGTCTTTATGTCTTGGGGTTGGGGTAGGTAGATACCAAATGAAACACTTTCAGGACCTT
-   : : : : : : : : : : : : : : : : : : : : : : : : : :
-   --GTCTTTACCTCT--GGGCTTGACTGGGCAGATGCCAGCTAAAAC---TTCCCTCACTGT
-   2780      2790      2800      2810      2820      2830

      3030      3040      3050      3060      3070
-   CCTTCCTCTTGCAGTTGTTCTTTAATCTCCTTTACTAGAGGAGATA-----AATATTTTG
-   : : : : : : : : : : : : : : : : : : : : : : : : : :
-   CTTTTCTATTTGACA---TCTTTCATCTCCTTTTCTATAGGTGACAGCAAGAATACTTTA
-   2840      2850      2860      2870      2880
```

Fig. 4E



13/23

```

3080      3090      3100      3110      3120      3130
CATATAATGAAGAAATTTTCTAGTATATAACGCAGGCCTTTTATTTTCTAAAATGATGA
::: :: : :: ::::: : :: :::: ::::::::::::::::::::
TATAGAACAAGGATATTTTTTT-----CAAGCCTGTTATTTTCTAAAATGA---
2890      2900      2910      2920      2930

3140      3150      3160      3170      3180      3190
TAGTATAAAAATGTTAGGATAACAGAATGATTTTAGATTTTCCAGAGAATATTATAAAGT
::: : :::: ::::: ::::: : ::::: : : ::::: : : ::::: :::::
TAGCACAAAC----TAGGACAACAGGATGATTTTCAGGTTTCTATATAAT-TTATAAAGT
2940      2950      2960      2970      2980

3200      3210      3220      3230
GCTTTAGGTATGAAAATAAATCATCTTTGTCTGATT
:::: : :::: ::::: ::::: : ::::: :
GCTTTGGATATCCAAATAAATCACCTTTGTCTGAGT
2990      3000      3010      3020

```

Fig. 4F

50